



## Geographical downscaling of outputs provided by an economic farm model calibrated at the regional level

P. Cantelaube, P.A. Jayet, Florence Carre, C. Bamps, P. Zakharov

### ► To cite this version:

P. Cantelaube, P.A. Jayet, Florence Carre, C. Bamps, P. Zakharov. Geographical downscaling of outputs provided by an economic farm model calibrated at the regional level. Land Use Policy, 2012, 29 (1), pp.35-44. 10.1016/j.landusepol.2011.05.002 . ineris-00961771

**HAL Id: ineris-00961771**

**<https://hal-ineris.archives-ouvertes.fr/ineris-00961771>**

Submitted on 4 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Geographical downscaling of outputs provided by an economic farm model calibrated at the regional level**

There is a strong need for accurate and spatially referenced information regarding policy making and model linkage. This need has been expressed by land users, and policy and decision makers in order to estimate both spatially and locally the impacts of European policy (like the Common Agricultural Policy) and/or global changes on farm-groups. These entities are defined according to variables such as altitude, economic size and type of farming (referring to land uses). European farm-groups are provided through the Farm Accountancy Data Network (FADN) as statistical information delivered at regional level. The aim of the study is to map locally farm-group probabilities within each region. The mapping of the farm-groups is done in two steps: 1/ by mapping locally the co-variables associated to the farm-groups, i.e. altitude and land uses; 2/ by using regional FADN data as *a priori* knowledge for transforming land uses and altitude information into farm-groups location probabilities within each region. The downscaling process focuses on the land use mapping since land use data are originally point information located every 18 km. Interpolation of land use data is done at 100 m by using co-variables like land cover, altitude, climate and soil data which are continuous layers usually provided at fine resolution. Once the farm-groups are mapped, European Policy and global changes scenarios are run through an agro-economic model for assessing environmental impacts locally.

### **Key-words:**

Downscaling, land use, spatial statistics, farm-groups, Farm Accountancy Data Network

## **1. Introduction**

There is a strong need for accurate and spatially referenced information regarding policy making and model linkage. This need has been expressed by land users, and policy and decision makers in order to estimate spatially and locally the impacts of European policy (like the Common Agricultural Policy) and/or global changes on economic agents and consequently on natural resources.

Economic models are able to provide information on changes of economic agents' behaviour according to different Policy scenarios whereas agronomic or environmental models can provide information on changes in natural resources according to global environmental changes. Within the Common Agricultural Policy (CAP), the farm-types are considered as being the economic agents. However, due to confidentiality issue, spatial data on farm-types are usually delivered at regional level, conversely to environmental data which are more freely accessible and usually available at local level. Then, modelling the local impacts of

European policy and /or global changes on economic agents and natural resources means:  
1/ to map locally farm-types dealt with economic models, 2/ to assess the impacts of the Policy on the spatial distribution of farm-types by linking economic models with agro-environmental models.

Locally mapping the farm-types from a regional level (small scale or coarse resolution) to a local level (large scale or fine resolution) is a downscaling issue or a disaggregation process (Bierkens et al., 2000) tackled as the core issue of the paper. The downscaling issue is tackled by using environmental information which can explain farm-types and which is usually provided at fine resolution and at European level. These are: land use taken as farm-type indicator, land cover, soil, climate, and topography used as co-variables. Two examples of application are presented: 1/ the assessment of CAP option scenarios on land use; 2/ the assessment of CAP option scenarios on agricultural N-pollutants.

Conversely to most of approaches used in spatial econometrics where spatial autocorrelation and spatial heterogeneity are basic characteristics of the models themselves, the downscaling occurs independently from the run of the economic model. In other words, we propose to downscale model outputs without changing a line of code into the model programming.

Vidal et al. (2001) noticed that changing the spatial unit used in statistical data redistribution is faster and more effective than setting up a new tool requesting new data. In addition we consider that changing the spatial distribution of data provided by a model from one geographical scale to another is faster than building a new model working at this new spatial unit. This is faster because this approach does not need to calibrate and to validate new models (we mention “models” at plural, taking into account the different fields and the large range of environmental problems for which the economic model could be used). This is more effective, regarding the amount and the quality of data called for a new model based on new spatial units and scale.

Summarily, downscaling of large sets of outputs provided by any run of the economic model through the one-step preliminary disaggregation procedure offers fast user-friendly means devoted to policy analysis and to model linkage. When the models and geo-referenced data are not hard linked, any update of one of them can improve the quality of the global modeling chain.

A first section is dedicated to the material used for the downscaling process where the general concept of the AROPAj economic model used is roughly explained and the economic and environmental data are described. A second section is devoted to the methodology, a

third one to the results, followed by a fourth section presenting the impact assessment applications.

## 2. Materials

### 2.1 General concepts of the AROPAj economic model

The AROPAj model is an economic model dedicated to the simulation of European Union (EU) farming systems' behavior facing change in economics (Common Agricultural Policy, European Directives regarding environment, the climate change and bio-energy). The AROPAj model aims at providing assessment of the impacts of European agricultural and environmental policies at different scales from the farm level to the EU one. Light presentations of the model are available in papers by De Cara and Jayet (2000), De Cara et al (2005). It is a short-term supply model using mathematical programming tools.

The main input of the AROPAj model is the Farm Accountancy Data Network (FADN<sup>1</sup>) in which samples supply economic information related to thousands of farmers across the European Union. In practice, estimates can be published only if derived from a sample of at least 15 farms, and the geographical identification consistent with representative information does not refer to a level finer than the "region" level.

The construction steps of AROPAj are (i) the farming design describing the farmer's objective (i.e. maximizing the gross margin), the activities and the bounded "production set" defined by constraints related to different blocks (i.e. crop systems, bovine demography, animal feeding, greenhouse gas emissions, nitrogen balance, ...); (ii) the classification of FADN samples into farm-groups (hereafter labeled as *farm-groups*) which are then the entities to be mapped at fine resolution (this step is described more in detail in section 2.2); (iii) pre-estimating of parameters through linear econometric methods calling for FADN data combined with expert information; (iv) calibration (i.e. re-estimating) of subsets of parameters through Monte Carlo and gradient methods. Different versions of the model are operational. In the next application section, we used the version covering the EU-15 and related to FADN-2002 (101 FADN regions are partitioned into a total of 1074 farm-groups). Each farm-group is assumed to be an individual economic agent, related to a maximization program. Solving of these programs calls for mathematical programming software combining linear programming and mixed integer variables (LMIP). This version represents a large part of "real" farms (a little less than 2 millions farms represented by the model) and of the utilized agricultural area (UAA) devoted to crops and pasture (UAA close to 88 million hectares represented by the model).

---

<sup>1</sup> [http://ec.europa.eu/agriculture/rca/index\\_en.cfm](http://ec.europa.eu/agriculture/rca/index_en.cfm). FADN is still the only representative and reproducible source of micro-economic data covering EU farming systems, even if costs and goals are widely questionable and discussed (Vrolijk et al, 2004).

Regarding the land use side, it has to be kept in mind that each individual LMIP delivers, among other variables, the optimal annual allocation of the individual UAA. This individual UAA is a fixed resource parameter. Any change in prices, in CAP or in the global environment (when the climate change is assumed to impact the crop growth function provided by the crop model) may lead to change in the land allocation at the farm-group level. Through the linkage to an agronomic crop model (i.e. STICS, see Brisson et al, 2002) AROPAj allows to simulate a wide range of alternative economic policy scenarios combined with various management and agro-environmental conditions.

## **2.2. Definition of the AROPAj farm-groups**

The second step of the AROPAj economic model leads to transform FADN regional data into EU representative farm-groups through hierarchical clustering applied to observed farms within each FADN region according to three criteria, (1) the average altitude (by 3 elevation classes, i.e. <300 m, 300-600 m and > 600 m), (2) the type of farming (defined by the FADN and disaggregated into 14 types of farming activities from which we exclude grape production and arboriculture), (3) the economic size unit (based on the total standard gross margin, which is defined by the value of output from one hectare or from one animal less the cost of variable inputs required to produce that output). The legal use of such data notably leads to restrain the number of clusters. Moreover agreement with the European Commission is required for an access to the FADN database.

## **2.3. Data used for mapping the farm-groups**

### **2.3.1 Altitude information**

The Digital Elevation Model (DEM) used is the *100mx100m STRM 2003 DEM*, a resample (and clipped on Europe) version of the digital topographic data from the SRTM project (NASA Shuttle Radar Topography Mission<sup>2</sup>) available from the Joint Research Center. The original 90m SRTM was resampled in accordance with the recommended INSPIRE guidelines (Commission of the European Communities, 2007): resampled at 100m, aligned with the INSPIRE reference grid and projected in the ETRS89 Lambert Azimuthal Equal Area. The altitude is categorized into 3 classes (less than 300m, between 300 and 600m, over 600m). Slopes were also computed and used together with altitudes in the land cover model (also categorized into 3 classes).

### **2.3.2 Available Land Use information at European level: LUCAS**

---

<sup>2</sup> <http://www2.jpl.nasa.gov/srtm>

LUCAS is the Land Use/Cover Area frame Statistical Survey (Delincé, 2001, Bettio et al, 2002, Gallego, ed., 2002<sup>3</sup>). This survey is coordinated by Eurostat and is based on direct observations described by a common typology all over Europe. The report of accurate and detailed geo-referenced information follows a detailed nomenclature, notably for agricultural categories: 57 land classes including 34 agricultural classes. The LUCAS is a non stratified systematic survey; it is delivered as point information sampled across the EU territory. The point sample consists of 10-points clusters which are set every 18 km (approximately 98 000 points are related to the EU15).

### 2.3.3 The land cover database: Corine Land Cover

As a part of the European CORINE Database, the CORINE Land Cover 2000 (CLC<sup>4</sup>) geographic database provides geo-referenced information on land cover over most European countries (including the EU 27 members states) (CEC-EEA, 1993, EEA 2001, JRC-EEA 2005). Accurate required data are accessible through the European Environmental Agency (EEA). It has been produced by photo-interpretation of satellite pictures (Landsat ETM images) and refined with ancillary data (aerial photographs, topographic maps, local knowledge ...). The CLC describes the land cover according to a 3 level nomenclature (5 classes for the first level – *Artificial surfaces, Agriculture, Forest, Wetland and Water Bodies* – and up to 44 classes for the more detailed level). CLC map has a scale of 1:100,000 (meaning that the location precision of CLC is 100m) and has a minimum mapping unit (the CLC polygons) of at least 25ha. The original vector land cover database converted into a raster format is made available at a 100mx100m raster cell size in Lambert-Azimuth coordinates (see EUROSTAT/GISCO Database Manual, Eurostat 2005). The CLC nomenclature is not detailed enough to distinguish between the different kinds of agricultural land cover. Arable crops are indeed aggregated (even for the 44 class nomenclature) and the CLC is a pluri-annual land cover map, consequently most crops are not reported (Gallego and Bamps, 2008).

### 2.3.4 The European Soil Database

The European Soil Database v2.0 (ESDB<sup>5</sup>) contains a large number of soil related parameters in raster data files with cell sizes of 1km x 1km. Those rasters are in the public

---

<sup>3</sup> *Area frame surveys: aim, principals and operational survey*, <http://agrienv.jrc.it/publications/ECpubs/agri-ind/> LUCAS has been deeply restructured in 2006 (Jacques and Gallego, 2005), but the data of 2001 were used in this study to “fit” with CLC 2000.

<sup>4</sup> CORINE Land Cover stands for CO-ordination on INformation of the Environment Land Cover See [http://www.ec-gis.org/docs/F10418/CLCTECHNICAL\\_GUIDE.PDF](http://www.ec-gis.org/docs/F10418/CLCTECHNICAL_GUIDE.PDF) and <http://ec.europa.eu/agriculture/publi/landscape/about.htm>

<sup>5</sup> [http://eusoils.jrc.it/esbn/Esbn\\_overview.html](http://eusoils.jrc.it/esbn/Esbn_overview.html)

domain but data access should be asked to the Joint Research Center and allow expert users to use the data, for instance to run soil, water and air related models/pedo-transfer rules. Parameters used in this study (all qualitative variables) are :

- Main and secondary limitation to agricultural use, dominant and secondary surface textural class, depth class of an obstacle to roots (from the sub-database Soil Geographic Database of Eurasia, SGDBE)
- Topsoil and subsoil water availability, topsoil and subsoil cation exchange capacity (from the sub-database Pedo-Transfer Rules Database, which holds results of application of pedo-transfer rules to SGDBE parameters).

These variables are quantitative.

### 2.3.5 The meteorological data

Meteorological data come from the European MARS meteorological database (Monitoring Agriculture with Remote Sensing<sup>6</sup>, MARS FOOD unit of the EU-JRC). Derived from global atmospheric model, this database holds daily climatic data for Europe interpolated on a 50 x 50 km grid. Data are daily, 10-daily and monthly outputs from the ECMWF (European Centre for Medium-Range Weather Forecast) atmospheric model. The two parameters extracted from this database to be used in this study are the *annual average temperature* and *annual sum of rain* (both parameters averaged on years 1990-2000). They are used as quantitative variables.

## 3. Method for downscaling the farm-groups

The starting point is assumed to be the availability of results provided by an economic model based on a set of representative farming systems. Due to lack of information or to limited access to information, geo-referenced modeled systems are usually not available at fine resolution. Nevertheless downscaling and mapping of such results can be achieved by using probabilistic linkage between economic activities and land cover.

Let us consider on the one side economic entities – namely farm-groups – defined according to variables such as altitude, economic size and type of farming. A realistic case occurs when statistical information is delivered at regional level. Let us consider on the other side that land cover information meets agricultural activities – or clusters of agricultural activities – accurately in terms of geographical location. These two information sources can be brought together, leading to assess the probabilistic location of economic entities at the finest level allowed by the land cover database.

---

<sup>6</sup> <http://agrifish.jrc.it/marsstat/datadistribution/> <http://agrifish.jrc.it/marsfood/ecmwf.htm>

As described in the material section, the very preliminary step simply consists in considering the agricultural supply-side model – namely the AROPAj model – which refers to a set of representative farm-groups defined at the regional level and acting as autonomous economic agents.

The next two steps related to mapping and downscaling issues, presented in the following, consist in assessing (i) probabilistic location of activities related to observable explanatory variables on a fine resolution grid, (ii) weights related to farm-groups at grid cell level. Concretely the method is implemented (i) by mapping locally the co-variables associated to the farm-groups, i.e. altitude and land use; (ii) by using regional data related to the Farm Accountancy Data Network (FADN) as *a priori* knowledge for transforming land use, land cover, climate, soil and altitude information into farm-group location probabilities within each region.

### **3.1 Combining the data into a common database**

The overlay relies on two GIS operations: (i) the overlay of CLC polygons with the other geo-referenced explanatory variables and also with the FADN regional boundaries; (ii) the combination of the LUCAS observations with layers of explanatory variables. The CLC is available in both a vector and a raster based format. The original vector land cover database converted into a raster format is made available at a 100mx100m raster cell size in Lambert-Azimuth coordinates (see EUROSTAT/GISCO Database Manual, Eurostat 2005). The overlay is realized at the regional level.

Co-location inaccuracies which are possibly accumulated during the different steps of the overlay are not discussed in this paper: We assumed that related errors refer to geographical deviations lower than 100m (see Gallego, 2002, for discussion and remarks about the CLC and LUCAS overlay).

The spatial unit resulting from this geographical overlay corresponds either to a CLC polygon (thus mapping unit of at least 25ha), either to an intersection of a CLC polygon with one (or more) of the other layer's mapping unit (ESDB mapping unit, cell from the 50x50m meteorological grid, DEM polygon or FADN region).

These new polygons are the mapping units of the geographic layer resulting from the downscaling process.

### **3.2 The downscaling process**

#### **3.2.1 General concept**

The downscaling consists in mapping the AROPAj farm-groups from the administrative regional FADN level (source unit) into a finer resolution level. One way to carry out such a



task consists in using co-variables and distributing data at the target level following the distribution of this co-variable at this target level, which is obviously assumed to be known (Vidal et al, 2001).

Co-variable information refers to soil, climate and elevation characteristics. Geo-referenced farm activities come through databases devoted to land use and land cover. It is assumed that these geo-referenced databases can be overlaid with FADN regions maps. Hazeu et al. (2010) built a methodology for spatially allocating farm-types. The procedure combines a logit model with a Bayesian highest posterior density estimator in order to distribute spatially an intermediate entity, the Homogenous Spatial Mapping Unit, defined by homogeneous production conditions (given by a combination of the co-variables) rather than administrative boundaries. Afterwards the farm-types are related to Homogenous Spatial Mapping Units (co-variable combinations). However, focusing the disaggregation on the combined intermediate entity rather than on each co-variable can lead to more errors difficult to identify (Bierkens et al., 2000).

The co-variable relates to the data through a mathematical function or a statistical model. We assume that the target unit is a partition of the source unit. Let us consider the quantitative variable  $Y$  located on the geographic unit  $A$  related to the source level. Let us consider a partition  $B_i$  of  $A$ . Let us define the variables  $Y_i$  which refer to the unknown distribution of  $Y$  over  $B_i$ . We set:

$$Y = \sum_{B_i \subset A} Y_i \quad (\text{Equation 1})$$

Let us use the set of co-variables  $X_i$  assumed to be known for subunits  $B_i$  with  $X$  “linked”/“related” to  $Y$ :

$$Y_i = f(X_i) \quad (\text{Equation 2})$$

$Y$  refers to an output related to a farm-group,  $A$  refers to a FADN region, and  $B_i$  refers to any sub-regional geographical level. Our problem can be so expressed: which would be the probability of locating a farm-group on any  $B_i$ . In other words, what is the “weight” of the considered farm-group in the agricultural activity on any  $B_i$ . A preliminary step consists in using the *Altitude*. This is obviously not sufficient for precise locating of the farm-group but taking into account this criterion will improve the likely location within the region.

Hence, as a first step of the GIS operation, the geographic unit  $A$  (to which the considered farm-group belongs) is restricted to the sub-units  $A_{alt}$  where *alt* refers to the elevation (according to the regions).

$$A = \bigcup_{alt=1,2,3} A_{alt} \quad (\text{Equation 3})$$

Regarding the co-variables required by the spatial disaggregation, the linkage between farm-groups and geo-referenced indicators of activities comes through the area devoted to agricultural land uses by farm-groups. Estimated areas are provided by the pre-estimating step of the AROPAj build-up (see the section 2.1). The generic variable  $Y$  is now replaced by a vector of areas  $S_{j,k}$  where  $j$  refers to the crop and  $k$  refers to the farm-group.

Finally the spatial disaggregation enters in a two-steps process (regarding the FADN region scale). The first step is devoted to the likely location of crops. The second step consists in estimating the likely location of regional farm-groups through the crop location. The first step is made of two sub-steps: (i) spatial econometrics is called for estimating the  $f$ -function when  $X$  refers to geo-referenced physical observations leading us to prior estimates of crop location; (ii) these prior estimates are used to assess the probability of crop location based on a cross entropy method.

### 3.2.2 Mapping the crop cover probability

Chakir (2009) proposed a 2 step-approach aiming at the estimate of the probability to find a land cover category (agricultural activity) on any geo-referenced cell. The first step is based on the application of a Multinomial Logit model (MNL) which relates the land cover/type of crops through the LUCAS points to a set of explanatory variables, within each FADN-region. As explained above, explanatory variables are CLC classes<sup>7</sup>, Altitude, Slope, Climatic parameters (average annual temperature and annual sum of rain), and Soil characteristic parameters. The MNL is a usual tool of prediction of the probability of occurrence related to a random event which here is the land cover category. The probability of finding the land cover category  $j$  on the cell  $i$  follows the expression:

$$P_{ij} = \frac{\exp(x_{ij}\beta_j)}{\sum_{j=1}^J \exp(x_{ij}\beta_j)} + \eta_{ij} \quad (\text{Equation 4})$$

in which  $x_{ij}$  denotes the explanatory variables,  $\eta_{ij}$  denotes the error term and  $\beta_j$  denotes the vector of parameters. In our problem the observed  $P_{ij}$  are provided by the LUCAS database.

The estimate of  $\beta_j$  leads to the estimate of the probability  $P_{ij}$  (denoted by  $\tilde{P}_{ij}$ ). The estimating process is conducted separately for each FADN region. Consequently the econometric model (i.e. the parameter  $\beta_j$ ) depends on the region.

---

<sup>7</sup> It is assumed here that CLC classes “Artificial surfaces” (mostly urban and industrial areas, numbered 1 in the level 1 nomenclature), “Water Bodies” and “Wetland” (nb.4 and 5 in the level 1 nomenclature) are not used for agricultural purpose; hence raster cells belonging to those classes are excluded from the modeling phase.

The second step aims at optimizing the estimates of the land cover probability. The previous econometric model does not take account of information provided by the FADN database such as the regional proportion of land cover related to the various crops. The land cover estimates are now refined in order to minimize the difference between the estimated land use share (derived from probabilities  $\tilde{P}_{i,j}$  estimated by the MLN model) and the observed land use share (derived from the FADN data). Chakir (2009) used the Generalized Cross Entropy method (GCE, see Rubinstein 1999, and Golan et al 1996) applied to the re-estimate of probabilities  $P_{i,j}$  of finding the crop  $j$  on the cell  $i$  (previously estimated by  $\tilde{P}_{i,j}$  and hereafter denoted by  $Q_{i,j}$ ). This leads to solve the following optimisation programme (separately for each FADN region) :

$$\min_{Q_{i,j}} \sum_i \sum_j Q_{i,j} \ln \frac{Q_{i,j}}{P_{i,j}} + \gamma \sum_j Q_j^E \ln Q_j^E$$

$$\text{subject to : } \forall i, j : Q_{i,j} \geq 0, \forall i : \sum_j Q_{i,j} = 1, \sum_j Q_j^E = 1, \forall j : \sum_i Q_{i,j} u_i = R_j + Q_j^E \varepsilon_j$$

in which  $R_j$  denotes the regional FADN area devoted to the land cover category  $j$  (including an additional category devoted to non-FADN land cover),  $u_i$  denotes the area of the  $i^{\text{th}}$  raster cell (i.e. the area of the polygons previously defined),  $\gamma$  denotes a user selected error weight, and  $\varepsilon_j$  denotes the components of a selected error vector (constrained by  $\sum \varepsilon_j = 0$ ). Obviously bias occurs regarding the fact that the FADN is based on yearly data (not synchronous with other databases) related to a sample of the agricultural holdings in the EU. Nevertheless we have refined estimates of land cover probability linked to FADN and useful for spatial disaggregation of the outputs provided by an agricultural economic model based on FADN. We denote the solution of the CGE programme by  $Q_{ij}^{LC}$ .

The whole procedure is summarized on the Figure 1.

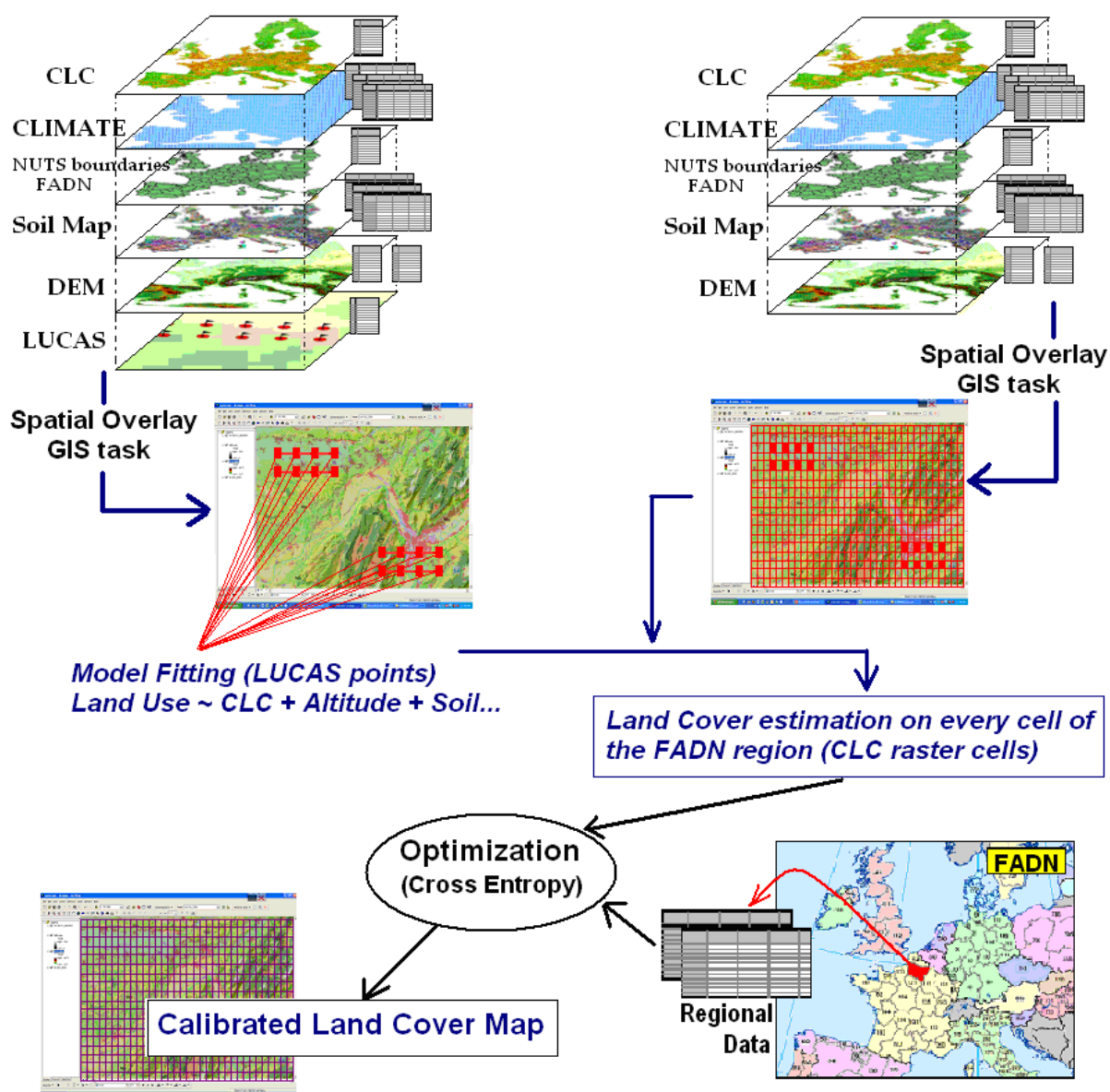


Figure 1 – Procedure to build a crop cover probability map from LUCAS, CLC2000 and other explanatory variables DEM, soil map and climatic data.

### 3.2.3 Mapping the farm-groups

Regarding now the co-variables required for the spatial disaggregation, the linkage between farm-groups and geo-referenced indicators of activities will come through the area devoted to the different agricultural activities (crop categories) for each farm-group. It is assumed that land use diversity within a region reflects the diversity among this region's types of farming. Estimated areas for each type of cultivation are provided by the pre-estimating step of the AROPAj build-up. Crop categorization into AROPAj distinguishes usual crop categories/land uses (Cereals, Permanent Crops, Permanent Grassland, Fresh Vegetables, Oleaginous, Fruit Trees, etc) but varies from one region to another due to the diversity in European

agriculture. For instance, category *Cereals* in some regions could be split into more classes : *Wheat*, *Maize* and *Other Cereals* (here other than *Wheat* and *Maize*), or *Wheat*, *Barley* and *Other Cereals* (now including *Maize* and not *Barley*), etc. On the contrary, in some Southern European regions, *Cereals* may be included into a general class *Other Crops* while other categories such as *Olive Trees* or *Vineyards* are taken into account.

Thus, for a given FADN region, let us note  $S_{jk}$  the surface occupied by crop-category  $j$  into farm-group  $k$ . The sum  $\sum_j S_{jk}$  provides the surface of each farm-group of the region while  $\sum_k S_{jk}$  provides the area occupied by each crop category into the region. The sum  $\sum_{jk} S_{jk}$  represents the total area (used for agricultural activities) of this FADN region.

Starting from a land cover probability map (at a fine resolution) distinguishing crop categories (as the ones used in AROPAj), a farm-group presence probability distribution can be derived from the crop category shares observed on the FADN region. Let us note  $Q_{ij}$  the probability of presence for crop category  $j$  on geographic unit  $i$ . Let us define  $T_{ik}$  related to the farm-group  $k$  and to the geographic unit  $i$ :

$$T_{ik} = \frac{\sum_{j=1}^J Q_{ij} S_{jk}}{\sum_{h=1}^K S_{jh}} \quad (\text{Equation 5})$$

Considering that all activities other than crops used in the first step are grouped in one extra land cover class ("Non Agricultural Use") and that the total number of land cover classes is  $J$ , it is easy to check that  $T_{ik}$  is eligible for a probability. For any FADN region we have :

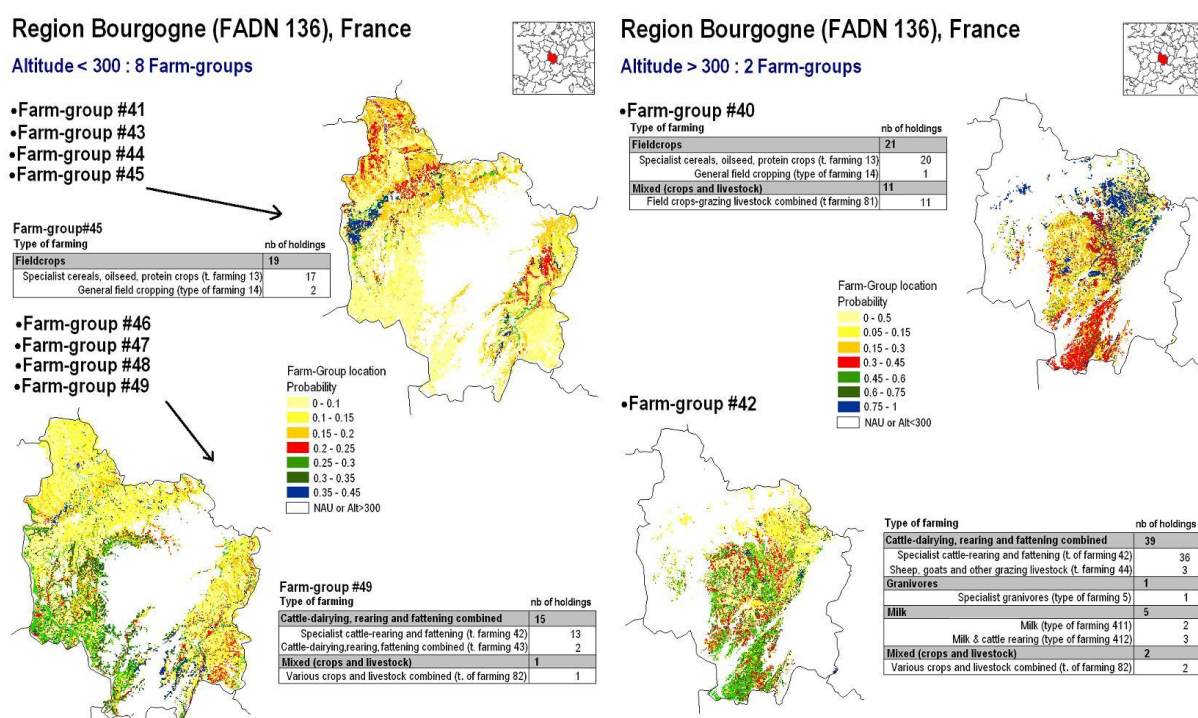
$\forall i, \forall k : 0 \leq T_{ik} \leq 1, \forall i : \sum_{k=1}^K T_{ik} = 1$ . The probability  $T_{ik}$  refers to the relative contribution of the  $k$  farm-group (within a region) to the share of agricultural activities present into the  $i$  cell belonging to the regional territory (with convenient altitude restriction). For a given  $k$  (in  $1, \dots, K$ ),  $\sum_i T_{ik}$  is equal to the area of the farm-group  $k$ .

Summarily (Equation 5) provides a probability distribution of the AROPAj farm-groups between the target units following the distribution of land cover categories on those target units. Let us recall that the  $f$  function stated by Equation (2) leads to the crop categories probability distribution  $Q_{ij}$  through the co-variable  $x$ . Availability of  $Q_{ij}$  allows us to implement a spatial distribution tool in order to scatter any output related to the farm-groups (that is to say any AROPAj output) over the FADN region.

## 4. Results of the downscaling

The spatial disaggregation of farm-groups is the result of the two-steps process described above. Following the first step devoted to the probabilistic location of crops (section 3.2.2), the second step consists then in estimating the probabilistic location of regional farm-groups through the crops location (section 3.2.3). The two steps have been implemented for numerous regions in the EU.

The results of probabilities of farm-group distributions are illustrated by the case of the Bourgogne French region. Figure 2 shows the probable location of farm-groups belonging either to the Altitude <300 class (right) or to the Alt>300 class (left). In the version of the AROPAJ model implemented on EU15 and related to FADN2002, the Bourgogne region leads to cluster FADN sample farms into 10 farm-groups ranged from “40” to “49” (8 on the Alt<300 side and 2 on the Alt>300 side). In addition the figure delivers the numbers of farms of the FADN sample split into FADN farm types which characterize a farm-group.



**Figure 2 - Probability of farm-group location in the Bourgogne region: 2 farm-groups (out of 8, #45 and #49) in altitude lower than 300 m, and the 2 farm-groups #40 and #42 in altitude higher than 300 m**

## 5. Applications of the downscaling

### 5.1 Assessment of CAP option impacts on land use

Back to the AROPAj model and to the AROPAj farm-groups, we use the downscaling tool to estimate the impacts of policy change on land use. Land allocation of farm-groups is seen as a typical output of the AROPAj model. The Common Agricultural Policy (CAP) offers a large range of applications which could impact the land allocation. Different CAP options are easily implemented into the AROPAj model thanks to the normative mathematical programming approach, and the model runs provide land allocation results for each farm-group and for any policy option.

We consider two CAP scenarios, namely the *Agenda 2000* and the *Luxembourg agreement*. The *Agenda 2000* refers to the CAP in effect before the reform adopted on June 2003 (i.e. the *Luxembourg agreement*). The de-coupling scheme regarding subsidy supply at the core of the *Luxembourg agreement* progressively enters in force in 2005 (see Jayet and Labonne, 2005, for theoretical and applied explanation regarding “de-coupling” in the CAP). Usual geographical visualization of the AROPAj outputs leads to monochrome mapping at the FADN region scale, as the aggregation of the farm-groups output. It is now possible to downscale the results thanks to the farm-groups’ contribution related to the agricultural activity spread over all 100mx100m cells within the FADN region. The downscaling process is applied to the mapping of change in the probability of presence of (i) pastures and (ii) cereals when the CAP turns the *Agenda 2000* scenario into the *Luxembourg agreement* scenario. Figure 3 and Figure 4 deliver maps which enlighten at the two levels (i.e. the region level on the left and the cell level for a few regions on the right) change in land allocation estimated as the difference of percentage of the total utilized agricultural area.

The figures underline clearly the dramatic change in the perception of CAP impacts regarding the resolution grid. Possible strong sub-regional variations can not be depicted when model outputs only refer to farm-group without fine mapping consideration.

Explanations of change come easily when we consider that CAP change, through support redirected from some inputs or products (direct payments devoted to crops and livestock) to others (the land), means strong change in real prices and correlatively change in land use. But change impacts are possibly quite complex. With CAP change, some plots (and farms) may be devoted to local higher yield crops while less numerous plots (or farms) are covered by these crops. That leads to possible relocation of production, and so reorganization of agricultural markets and logistics up-stream and down-stream. Some of other dramatic consequences of these changes arise regarding the environment, mainly through change in pollution sources (i.e. Nitrogen losses).

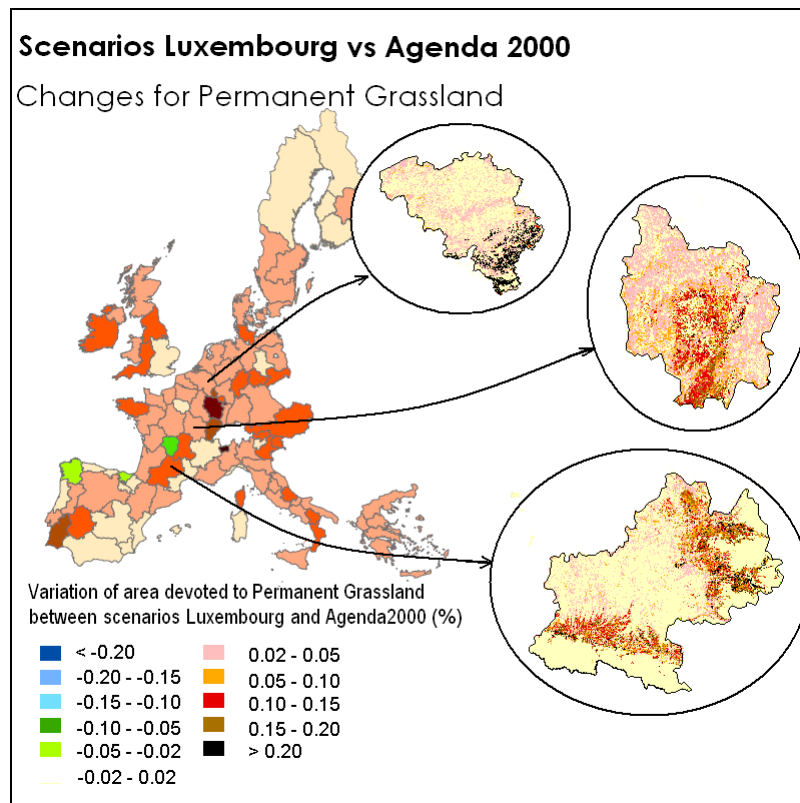


Figure 3 – Change in probability of presence of *Pasture* when the CAP turns the *Luxembourg agreement* scenario into the *Agenda2000* scenario: mapping at the regional level (EU15, left) and at the 100x100m raster cells level for the FADN regions *Belgium*, *Bourgogne* and *Midi-Pyrénées* (right).

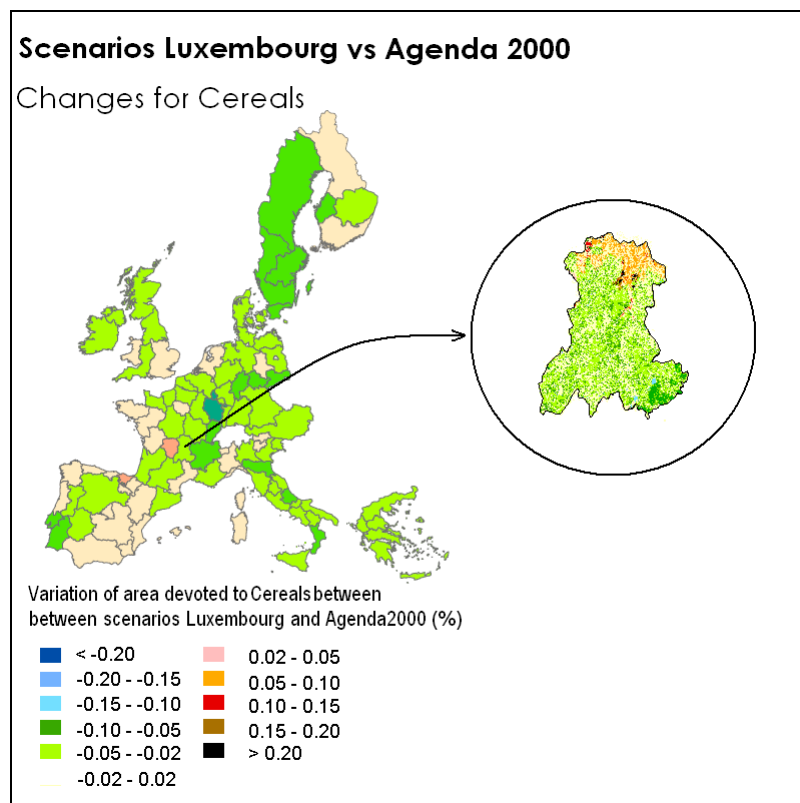


Figure 4 – Change in probability of presence of *cereals* when the CAP turns the *Luxembourg agreement* scenario into the *Agenda2000* scenario: mapping at the regional level (EU15, left) and at the 100x100m raster cells level for the region *Auvergne*.



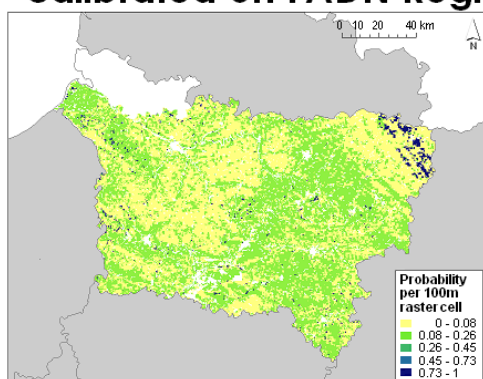
## 5.2 Spatial disaggregation of agricultural N-pollutants

Knowledge of accurate location of pollution is of high interest regarding at least two aspects that are (i) display of information useful for policy makers and (ii) disposal of quantitative elements as a part of a complete chain of modelling. The latter occurs when an agro-economic model is linked to a hydro-geological model dedicated to the estimate of nitrate concentration in aquifers. More precisely, the AROPAj model will be linked to the hydrological model MODCOU (Ledoux et al, 2007), which describes surface and groundwater flow at a daily time step. Time and spatial disaggregation will be requested. We focus here on the spatial dimension, when AROPAj outputs need to match the MODCOU grid cells (squared cells from 1 to 16 km<sup>2</sup>). Our downscaling tool is applied to the AROPAj estimates of yearly presence of N-pollutants sourced from agriculture at the convenient geographical scale.

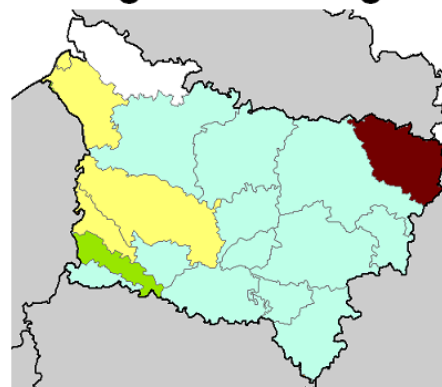
The AROPAj model has been strongly improved by introducing N-yields functions based on the use of the STICS crop model. The adequate methodology is developed by Godard et al (2008). Additional work devoted to greenhouse gas emission by Durandea et al (2010) has been extended to provide N-pollution functions regarding nitrous oxide (N<sub>2</sub>O, a greenhouse gas), ammonia (NH<sub>3</sub>) and nitrate (NO<sub>3</sub>) mainly sourced from the use of nitrate fertilizers. We provided a set of functions applied to the French agriculture regarding the major crops. The N-yield functions are expressed by  $Y(N) = b - (b - a)e^{-\tau N}$  in which  $b$  denotes the asymptotic potential,  $a$  denotes the yield related to the 0 N-level, and  $\tau$  relates to the curvature, and N-pollutant functions are adjusted to expression of the type  $P(N) = c + d N$ . All parameters  $a$ ,  $b$ ,  $\tau$ ,  $c$ ,  $d$  depend on the crop and on the farm-group.

Example: Land cover category **Permanent Grassland**  
on FADN Region 132 ("Picardie", France)

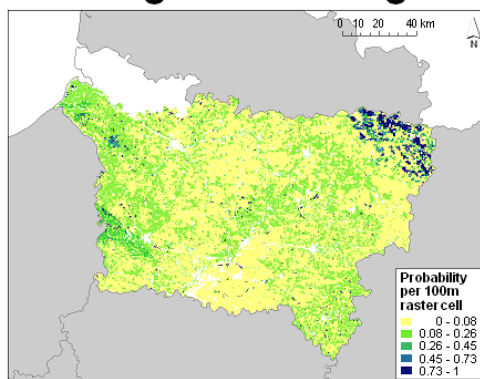
**① Land Cover map calibrated on FADN Region**



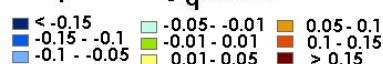
**② Evaluation on Small Agricultural Regions**



**③ Land Cover map calibrated on Small Agricultural Regions**

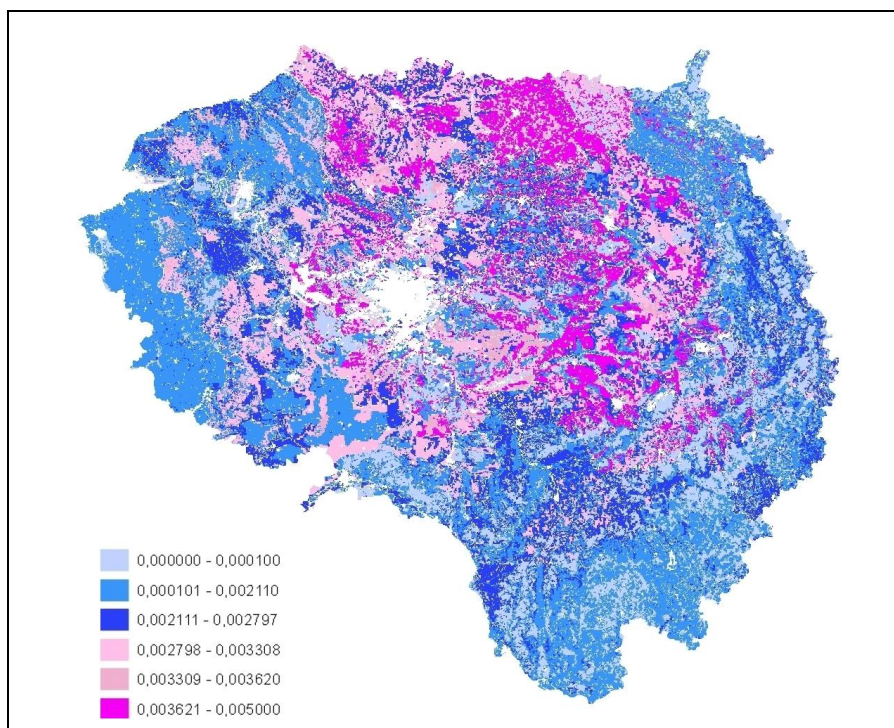


$q_{\text{Reference}} - q_{\text{Method}}$

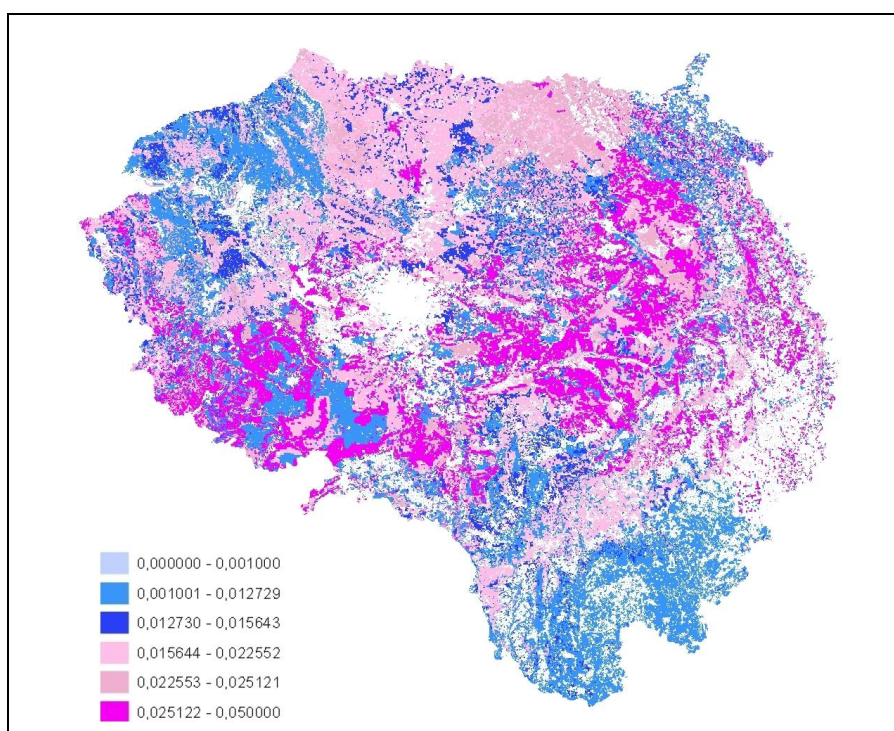


$q$  = share of the regional area occupied by Permanent Grassland between all the Small Agricultural Regions.

**Figure 5 – Procedure to refine the land cover map with data provided on the Small Agricultural Regions. Example for FADN region 132 ("Picardie", northern France)**



**Figure 6 – N-loss in N<sub>2</sub>O emissions (tN<sub>2</sub>O/ha) sourced from the agricultural activities represented by the AROPAj model on the Seine river Basin (Agenda 2000 scenario).**



**Figure 7 – N-loss in NO<sub>3</sub> emissions (tN<sub>NO3</sub>/ha) sourced from the agricultural activities represented by the AROPAj model on the Seine river Basin (Agenda 2000 scenario).**

In this subsection the study area is limited to the Seine river basin (Northern France) which intersects 8 FADN regions and 54 farm-groups, covering more than 9 millions hectares. We use additional information regarding the distribution of agricultural activities (areas of each

land cover category) available at a level finer than the FADN region. This finer level refers to small districts named Small Agricultural Regions (SAR)<sup>8</sup>. In order to take full advantage of this information, the land cover map is re-calibrated from the GCE step at the SAR district level instead of at the FADN region level as previously (see the 3.2.2 subsection). The farm-group location is now based on these new estimates of the probability  $Q_{ij}^{LC}$  and re-estimated through (equation 5). The process is illustrated by the Figure 5. The two land cover maps (numbered 1 and 3 on Figure 5) refer to the same total (regional) surface regarding each land cover category (permanent grassland on the figure). But change in the calibration may significantly impact the SAR surface related to a given land category (see the map 2 on Figure 5).

Lastly AROPAj runs lead to the estimate of farm-group outputs including losses of N in the various categories for which the STICS crop model is designed. The mapping of down scaled results is displayed on Figure 6 (N<sub>2</sub>O) and Figure 7 (NO<sub>3</sub>). It should be noticed that the results only refer to the N-loss related to agricultural activities represented by the AROPAj model. Artifacts remain in term of delimitation when administrative boundaries seem to overlay color change. They can probably be explained by the farm clustering leading to aggregation bias. In addition, the limited number of intervals (i.e. colors) partly explains them. And we can suspect other bias source when all databases used in the modeling chain differ in years. Nevertheless accuracy of the (likely) location of the economic model outputs will meet a large range of applications devoted to policy analysis and to model linkage.

## 6. Concluding remarks

Our downscaling method helps to achieve the spatial distribution of the AROPAj economic model outputs at the finest geographical resolution level inferred by the most geographically detailed database. This is strongly useful for land use policy making when impact heterogeneity enters the decision process as well as for coupling the economic model with physical models when the analyze of environmental phenomena requests information on high spatial resolution grids. In this sense, the mapping of impacts related to different agricultural policies is a matter of great concern.

The spatial distribution of outputs is provided through the localization of the economic working units of the model, i.e. the farm-groups. The mapping of AROPAj farm-groups is based on the land allocation among agricultural activities related to all farm-groups inside

---

<sup>8</sup> Small Agricultural Regions (in French '*Petites Régions Agricoles*'), units defined by a regrouping of communes within a '*Département*' (subregional partitioning) having 'similar' agricultural orientations and activities. It is managed by the French institute for statistics and economy (INSEE).

[http://agreste.agriculture.gouv.fr/definitions\\_6/zonages\\_81/index.html](http://agreste.agriculture.gouv.fr/definitions_6/zonages_81/index.html)

Those data were provided by PIREN-Seine project partner *INRA SAD-MIRECOURT*.

each FADN region. This statistical information in term of location probability is linked to land use information which is interpolated mainly by using Corine land cover as a support (this means that intensive animal producers would be worse located). CORINE land cover is the main support used to establish the land cover map and thus the spatial unit for the disaggregation is the 100x100m CLC raster cells.

The farm-groups location map allows us to highlight the diversity of CAP impacts on land use, as it is shown by the application of the method to assess the impacts of the Luxembourg reform implementation. It is clear that other outputs provided by the AROPAj model through the farm-groups mapping would modify the analysis of land use change from set-aside to abandonment, compared to what a more usual modeling approach based on homogenous regional or national scale would show. In the same line of thinking, the analysis of environmental indicators such as greenhouse gas emissions or such as nitrous pollutants would be strongly affected by the high resolution mapping of results, when policy makers have to take account of individual consequences of policies for making their implementation easier.

With respect to the existing right of access to databases and information, we design a route for the mapping of economic model outputs which is something else than fictive homogeneous regional or national maps. The increasing availability of richer information and the improvement of statistical methods related to spatial disaggregation joined with geographical information systems and more performing computation tools, lead to supply geographically disaggregated results. Two use fields appear. Firstly, a better assessment of environmental impacts related to agricultural activities requires high resolution mapping. This is useful for the improvement of the economic regulation of external effects as well as for the improvement of environmental impact assessment related to any change in policy and in the environment (e.g. climate change). Secondly, back to economic theory and to policy design, interesting perspectives appear regarding for instance the subsidiarity principle in policy making.

## **Acknowledgment**

The GENEDEC project was funded by the European Union under the contract 502184 (2004, 6<sup>th</sup> Framework Programme) and aimed at analysing the impacts of CAP change on agriculture and the environment.

The perennial French interdisciplinary PIREN-Seine programme aims at understanding the ecological and economical functioning of the Seine River basin.

We thank participants and the discussant during the contributed paper session of the XII<sup>th</sup> EAAE congress held in Ghent on August 2008 for helpful comments.

Technical data overlay at the beginning of the procedure was realized by the JRC team involved in the GENEDEC programme.

We strongly thank Delphine Leconte-Demarsy for her careful reading of the paper.

## References

- Bettio M., Delincé J., Bruyas P., Croi W., Eiden G., 2002, Area frame surveys: aim, principals and operational surveys. in Building Agri-environmental indicators, focussing on the European Area frame Survey LUCAS. EC report EUR 20521, pp. 12-27  
<http://www.ec-gis.org/docs/F22064/BUILDING%20AGROENV%20INDICATORS.PDF> (consulted the 5th of November 2010)
- Bierkens F.P., Finke P.A., de Willigen, P., (Eds) 2000. **Upscaling and Downscaling Methods for Environmental Research**. Chap. 3, Vol 88 "Developments in Plant and Soil Sciences" Series. Kluwer Academic Publishers.
- Brisson N., Ruget F., Gate P., Lorgeou, F and Nicoullaud B., 2002, **STICS: a Generic Model for the simulation of crops and their water and nitrogen balances, Model Validation for Wheat and Maize**, *Agronomy* 22, 69-92.
- CEC EEA, 1993, **CORINE Land Cover; guide technique**, Report EUR 12585EN. Office for Publications of the European Communities, Luxembourg. 137p.
- Commission of the European Communities (2007). **Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)**, Official Journal of the European Union L 108/1.
- Chakir R., 2009, **Spatial Downscaling of Agricultural Land–Use Data: An Econometric Approach Using Cross Entropy**, *Land Economics* vol. 85, pp. 238–251.
- De Cara, S., Houzé M., and Jayet P.-A., 2005, **Methane and nitrous oxide emissions from agriculture in the EU: A spatial assessment of sources and abatement costs**. *Environmental and Resource Economics* 32(4), 551-583
- De Cara, S. and Jayet, P.A., 2000, **Emissions of greenhouse gases from agriculture: the heterogeneity of abatement costs in France**. *European Review of Agricultural Economics*, 27(3): 281-303.
- Delincé J., 2001, **A European approach to area frame survey**. *Proceedings of the Conference on Agricultural and Environmental Statistical Applications in Rome (CAESAR), June 5-7, Vol. 2 pp. 463-472*  
[http://mars.jrc.it/mars/content/download/566/4117/file/Lucas\\_CAESAR.pdf](http://mars.jrc.it/mars/content/download/566/4117/file/Lucas_CAESAR.pdf) (consulted the 5th of November 2010)
- Durandea S., Gabrielle B., Godard S., Jayet P.A., Le Bas C., (2010), **Coupling biophysical and micro-economic models to assess nitrous oxide emissions from cropland and the effects of mitigation measures on greenhouse emissions**, *Climatic Change* vol. 98, n° 1–2, pp. 51-73.



EC-EEA, 1993. **CORINE Land Cover; technical guide**. Report EUR 12585EN. Office for Publications of the European Communities. Luxembourg,. 144 pp <http://www.eea.europa.eu/publications/COR0-landcover/page001.html> (Consulted the 5th of November 2010)

EEA, 2001, **Towards agri-environmental indicators; Integrating Statistical and administrative data with land cover information**. Topic Report 6/2001. Office for Publications of the European Communities, Luxembourg. 131p.

Elbersen B., Kempen, M., van Diepen K., Andersen E., Hazeu G., Verhoog D., 2006, **Protocols for spatial allocation of farm types**, SEAMLESS Report No.19, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, 107 pp, ISBN no. 90-8585-046-0.  
[www.seamless-ip.org/Reports/Report\\_19\\_PD4.7.1.pdf](http://www.seamless-ip.org/Reports/Report_19_PD4.7.1.pdf) (consulted the 5th of November 2010)

Hazeu G., Elbersen B., Andersen E., Baruth B., van Diepen K., Metzger M., 2010, **A biophysical typology in agri-environmental modelling**, chap.7 in Environmental and Agricultural Modelling, Brouwer F and van Ittersum M Eds, Springer, pp 159-187.

Gallego F.J., 2002, **Building Agro Environmental Indicators. Focusing on the European area frame survey LUCAS**. EUR Report 20521 EN. 176 p.

Gallego J. and Bamps C., 2008, **Using CORINE land cover and the point survey LUCAS for area estimation**, *International Journal of Applied Earth Observation and Geoinformation*, v. 10, 4, p. 467-475.

Godard C., Roger-Estrade J., Jayet P.A., Brisson N., Le Bas C., (2008), **Use of the available information at a European level to construct crop nitrogen response curves for the regions of the EU**, *Agricultural Systems*, vol. 97, n° 1–2, pp. 68–82. April 2008

Golan, A., Judge, G. and Miller, D.: 1996, **Maximum entropy econometrics: Robust estimation with limited data**, New York, Ed. John Wiley and Sons, New York

Jacques P., Gallego F.J., 2005, **The LUCAS project – The new methodology in the 2005/2006 surveys**. Workshop on Integrating agriculture and environment: CAP driven land use scenarios, Belgirate, September 26-27, 2005  
<http://mars.jrc.it/mars/content/download/567/4122/file/Lucas%20new%20methodology.pdf> (consulted the 5th of November 2010)

Jayet P.A., Labonne J., 2005, **Impact d'une réforme de la Politique Agricole Commune par le découplage**, *Economie et Prévision* n°167, 2005-1, pp 101-116.

JRC-EEA, 2005, **CORINE Land Cover updating for the year 2000: Image2000 and CLC2000; Products and methods**; ed. Vanda Lima, Report EUR 21757 EN. JRC-Ispra

Ledoux, E., Gomez E., Monget J.M., Viavattene C., Viennot P., Ducharne A. Benoit M., Mignolet C., Schott C., Mary B., 2007, , **Agriculture and groundwater nitrate contamination in the Seine basin – The STICS-MODCOU modelling chain**, *Science of the Total Environment*, Special issue, vol 375, n°1-3, pp 33-47, April 2007.

Panagos, P. , 2006, **The european soil database**. *GEO: Connexion*, 5(7), 32-33.

Rubinstein, R.Y, 1999, **The Cross-Entropy method for combinatorial and continuous optimization**, *Methodology and Computing in Applied Probability*, 2, 127-190.

Vidal, C., Gallego, F.J., Kayadjanian, M., 2001, **Geographical use of statistical data, methodological overview**. in “*Towards agri-environmental indicators*”, EEA Topic Report 6/2001, p 11-24.

Vrolijk, H. C. J., Meier, B., Kleinhanß, W., Poppe, K. J., 2004, **FADN: Buttress for farm policy or a resource for economic analysis?** *EuroChoices*, 3(3), 32-37.